# Joint multisensor exploitation for mine detection

Scott G. Beaven[*a], Alan D. Stocker[a], Edwin W. Winter[b]
[a]Space Computer Corporation, 12121 Wilshire Blvd, Suite 910, Los Angeles, CA 90025
[b]Technical Research Associates, Inc., San Diego, CA 92109

## ABSTRACT

Robust, timely, and remote detection of mines and minefields is central to both tactical and humanitarian demining efforts, yet remains elusive for single-sensor systems. Here we present an approach to jointly exploit multisensor data for detection of mines from remotely sensed imagery. LWIR, MWIR, laser, multispectral, and radar sensor have been applied individually to the mine detection and each has shown promise for supporting automated detection. However, none of these sources individually provides a full solution for automated mine detection under all expected mine, background and environmental conditions. Under support from Night Vision and Electronic Sensors Directorate (NVESD) we have developed an approach that, through joint exploitation of multiple sensors, improves detection performance over that achieved from a single sensor. In this paper we describe the joint exploitation method, which is based on fundamental detection theoretic principles, demonstrate the strength of the approach on imagery from minefields, and discuss extensions of the method to additional sensing modalities. The approach uses pre-threshold anomaly detector outputs to formulate accurate models for marginal and joint statistics across multiple detection or sensor features. This joint decision space is modeled and decision boundaries are computed from measured statistics. Since the approach adapts the decision criteria based on the measured statistics and no prior target training information is used, it provides a robust multi-algorithm or multisensor detection statistic. Results from the joint exploitation processing using two different imaging sensors over surface mines acquired by NVESD will be presented to illustrate the process. The potential of the approach to incorporate additional sensor sources, such as radar, multispectral and hyperspectral imagery is also illustrated.

**Keywords:** Mine detection, multi-sensor, robust detection, fusion

## 1. INTRODUCTION

Robust detection of mines and minefields over wide areas presents a significant challenge to both humanitarian and tactical demining efforts. Ground-based methods are effective, but are limited in terms of search rate and coverage area, leading to the need for airborne systems for wide-area mine detection. The objective of this effort was to establish a proof-of-concept for improved minefield detection through the joint exploitation of multiple sensors from airborne platforms. In this paper we establish the foundation for improving minefield detection performance through development of effective multi-sensor processing methods for detecting surface minefields from airborne sensors. This effort concentrates on the use of existing sensor data acquired by Night Vision Electronics and Sensors Directorate (NVESD) using mid-wave infrared (MWIR) and laser-illuminated imagery in the near infrared. Space Computer Corporation has developed methods for joint algorithm exploitation (JALEX) [Stein, et al. 2001], and adapted these for use in joint multisensor exploitaiton (JMEX). These approaches are based on the notion of fusion of multiple detector outputs to improve target sensitivity, while reducing false alarm rates over individual detection algorithms or sensing modalities. The method is based on sound detection theory principles as an alternative to *ad hoc* methods, such as simple Boolean logic. The approach uses pre-threshold detector outputs to formulate accurate models for marginal and joint statistics across multiple detection or sensor features. This joint decision space is modeled and decision boundaries are computed to provide a robust multi-algorithm or multisensor detection statistic. In this paper we describe application registration processing and the joint exploitation methodology for fusion of minefield discriminants from multiple sensor observables. The application of JMEX processing to multisensor airborne measurements over

minefields results in significant detection performance improvements over individual sensor processing results.

## 2. JOINT MULTISENSOR EXPLOITATION OVERVIEW

The quest for systematic performance enhancement from application of multiple sensors or algorithms to detection of minefields targets has led to the recent development of joint exploitation processing. The JMEX approach is a systematic methodology for fusing registered multiple detector outputs. The goal of this approach is to adaptively combine multiple detector outputs such that performance is improved (or at a minimum, retained) over the best single sensor performance. This methodology is based on development of a fully adaptive decision criterion from the joint distribution formed by the multiple detector outputs. Ideally we would like to define an optimal decision contour to segregate target from clutter in the joint decision space as illustrated in figure 1. The JMEX approach seeks to approximate an optimal decision boundary through a process that includes modeling of the marginal probability distributions of each sensor detection feature, standardizing these marginals and accounting for the correlation between them. Evaluation of the joint distribution for each sample results in a scalar fusion decision statistic such that is thresholded for detection purposes. Each threshold corresponds to a decision contour in the joint detection feature space from multiple sensor detectors.

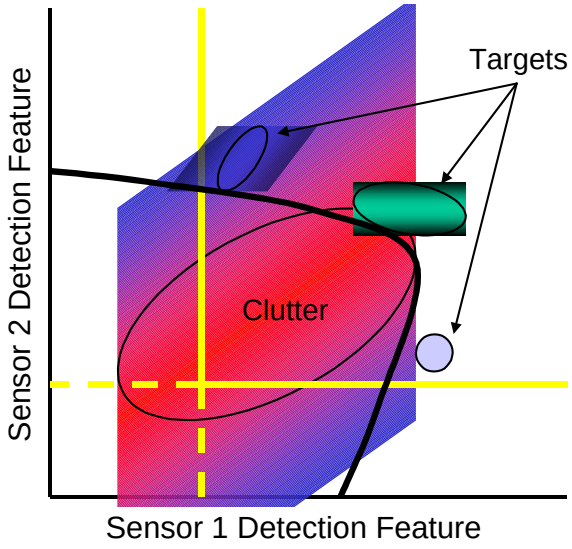



**Figure 1.** Conceptual illustration of joint fusion of detection features from two sensors.

The general framework for joint multisensor processing is illustrated in figure 2. This paper focuses on the use of multi-band MWIR and Laser-illuminated near-IR sensor data with SCC's registration and joint multiband, multisensor exploitation framework. This general fusion framework is flexible and can be expanded to accommodate other maturing sensor technologies for mine detection including scanned lasers, pulse compression lasers (LADARs), imaging radars and hyperspectral sensors operating in the reflective EO and thermal bands. The input data consists of multispectral MWIR and Laser-illuminated imagery, with potential future extensions to broadband and multi-band LWIR and Hyperspectral Imagery (HSI) data. The framework includes the capability for cross-sensor registration, detection feature generation (based on anomaly detection), joint processing, and post detection spatial/spectral analysis. For general mine detection application the detection features are based on anomaly detection filtering, since signatures for mines may vary widely and are not typically known *a priori*.

Our approach combines geolocation based registration with scene-based refinement to generate multisensor imagery that can be jointly exploited. Since registration accuracy from multiple sensors cannot be accomplished in general at the subpixel level, JMEX processing approach accommodates joint exploitation at the object level, based on detection features derived from each source. Detection features are computed from application of mature anomaly detection algorithms. The derived detection features are combined

optionally using our JALEX approach, which is described in detail later, and then fed into the JMEX processing to fuse features derived from multiple imaging sensors.

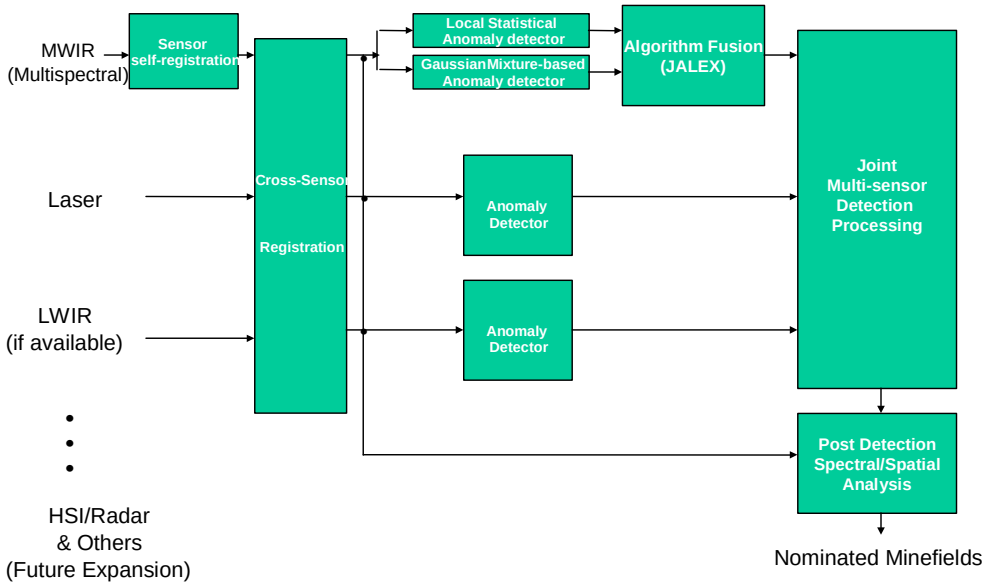



**Figure 2.** Joint Multisensor Exploitation Processing Framework.

The method is based on a fully adaptive decision criterion derived from the joint distribution formed by the multiple detector outputs. The JMEX approach seeks to approximate an optimal decision boundary through a process that includes modeling of the marginal probability distributions of each sensor detection feature, standardizing these marginals and accounting for the correlation between them. Evaluation of the joint distribution for each sample results in a scalar fusion decision statistic such that is thresholded for detection purposes. The result from JMEX processing is a detection statistic plane upon which we apply spatial clustering with optional spatial filtering and cluster size and shape constraints to produce individual detected object locations.

## 3. REGISTRATION and DATA PREPARATION

Two primary sensor data sources were made available by NVESD for this effort: 1) a mid-wave infrared (MWIR) and 2) a laser-illuminated near-infrared imager. The multisensor data set includes imagery from minefields in short grass, tall grass and primarily bare soil with sparse vegetation from a single-band MWIR imager and a laser-illuminated two-polarization imager operating at 808 nm. The sensors were deployed as part of the Lightweight Airborne Multispectral Minefield Detection (LAMD) program at NVESD. The laser system is an SMD BT-25 low-light Breadboard Laser camera that uses high-throughput dual polarizing beam splitters to provide P and S channel separation. The MWIR camera is an Amber Radiance 1 system with multiple filters to provide different band selections. The multisensor data set was used to refine georegistration and image-based registration techniques and for evaluation of our multisensor exploitation processing techniques.

### *Multi-Frame/Multi-Sensor Registration*

The multisensor data set described above was used to perform cross-sensor registration, as well as to address the issue of sensor self-registration from frame-to-frame. These data, along with corner point geoposition information, were used to produce georegistered image mosaics suitable for multisensor exploitation. There were several key issues that we resolved in generating these results. First, the imagery

(particularly laser camera) appeared to contain a significant amount of residual non-uniformity. We minimized this artifact by applying a scene-based gain correction to the data. This scene-based information was carefully derived to avoid the effect of a few dominant scatterer in the imagery (such as fiducial panels). Second, there were residual frame-to-frame geopositions or jitter errors that had to be resolved to register subsequent overlapping frames. These errors were small in absolute terms (e.g. 5-15 cm) but resulted in significant degradation of imagery when multiple frame data were averaged to produce the georeferenced image product. Finally, residual-scaling factors from the "jitter"-corrected imagery had to be resolved between the sensors. This scaling was observed to be < 2%, and is not apparent in viewing the individual frame data, but becomes apparent when a large number 50-60 frames are projected onto a geo-referenced grid. These were resolved by estimating this minor scaling difference from the imagery.

***Residual Displacement Correction of Georeferenced Imagery***

Initial attempts to produce a mosaicked set of imagery from the series of partially overlapping georeferenced image frames resulted in apparent blurring. This blurring was due to small geo-positioning errors and/or jitter in the range of a few pixels (5-15 cm). The source of the error may either be due to location accuracy of GPS and point information or imprecision in camera models used to derive image corner point geolocation information. This small error source is particularly apparent when forming a mosaic from multiple overlapping frames of image data. The left side of figures 3 shows the result from averaging a number (typically 2-5 for this data set) of overlapping frames for both the MWIR imager. These images over panels in the scene illustrate that there is a small, but significant offset between geo-registered image frames once they are projected into geolocation-based coordinates. The result of this is a blurring of the imagery and in the case of the panels, ghosting of localized objects. This residual error varies within the data set but typically ranges from 0 to 15cm.

To correct this residual frame-to-frame misregistration we incorporated local-area-based estimation of frame-to-frame displacement ("jitter-correction") for overlapping regions of adjacent frames. These estimates were used to define a relative offset for each frame of data and resulted in a more precisely registered image product. The right sides of Figures 3 illustrates the resulting improved composite georeferenced imagery. The processing described above was used to generate sets of geo-referenced, registered image mosaics that are suitable for cross-sensor analysis and processing. The resulting sensor-to-sensor registration is sufficient for object-level multisensor analysis and processing as we initially proposed. Examples from both laser-illuminated and MWIR imagery are shown in Figure 4. In both cases we show the georeferenced image mosaic on the left, with detail highlights of a selected region on the middle and right of each figure.

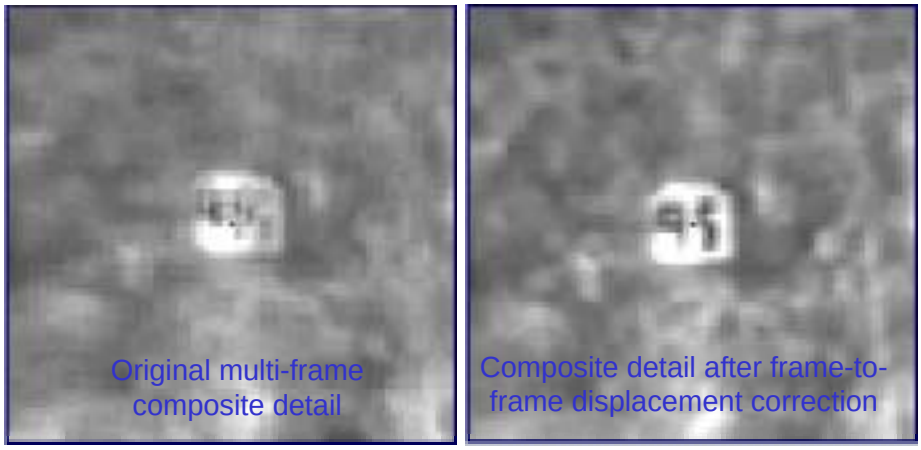



Figure 3. Multiframe composite MWIR image detail without (left) and with (right) frame-to-frame jitter-correction. Note the significant improvement in image clarity on both the panel and natural background within this scene. Frame-to-frame displacement correction is incorporated into our multisensor geo-referenced imagery processing.

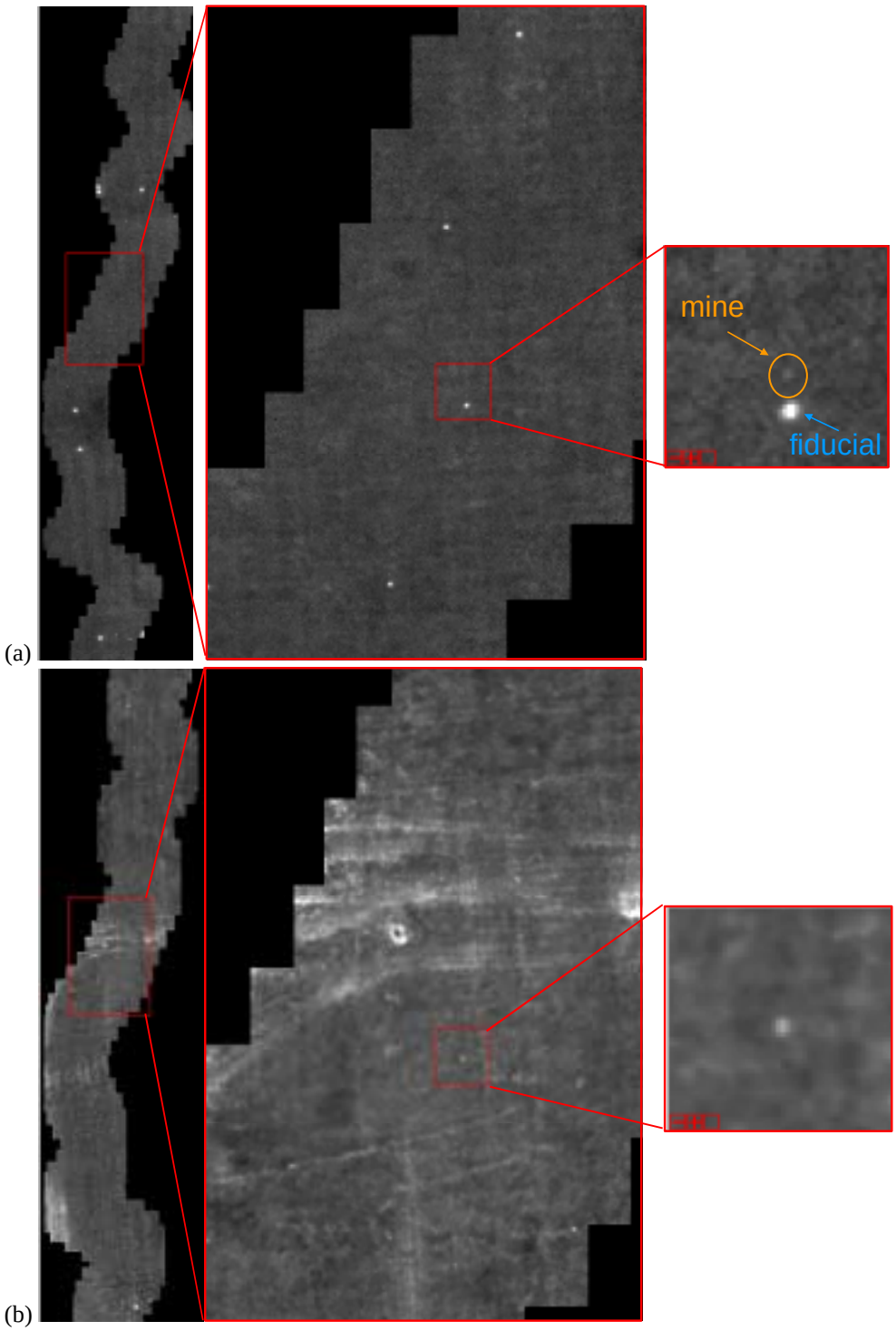



Figure 4. a) Georeferenced display of Laser-illuminated (co-polarized) imagery. This includes the residual non-uniformity and displacement (jitter) correction, is well as dynamic multi-frame averaging to produce this georeferenced image mosaic. b) Georeferenced display of MWIR imagery. This includes the residual non-uniformity and displacement ("jitter") correction, as well as dynamic multi-frame averaging to produce this georeferenced image product.

## 4. DETECTION FEATURE GENERATION

The registration processing described above was used to generate 3-band image cubes containing the two laser bands and the single MWIR band to examine efficacy of joint exploitation of multisensor processing. Figure 5 shows the three bands of the georegistered imagery from a section run over tall grass used in our analysis. We reduced the sampling of both sensors to facilitate data handling and further processing on these data. Because of the residual registration error between sensors, processing of the multisensor imagery at the pixel level is not feasible. Therefore we applied the capability in our JMEX approach to associate objects from individual detection feature planes based on relative positions. This simple association approach relaxes the need for precision location, but still requires that objects be within a reasonable distance from each other to ensure adequate multisensor detection performance.

To generate individual sensor detection features we applied a local Reed Xiaoli (R-X) anomaly detector implementation [Reed and Yu, 1990; Haskett and Rupp, 2002]. We adjusted the algorithm to optimize performance over each of the scenes processed. Although we attempted other anomaly detector approaches, found that the local R-X algorithm consistently worked well over the imagery analyzed and was selected to be the initial detection feature used for joint processing. Since the clutter in the scenes used was relatively benign, we did not observe any advantage in the segmentation that spectral clustering or linear mixture-based algorithms. We believe that more severe clutter environments may require the use of these other approaches for detection feature extraction and the framework we are employing here will accommodate this. The R-X algorithm was applied to the two-band laser imagery independent of the MWIR image. Figure 5 shows a section of the MWIR imagery with the raw R-X filter output. The corresponding result from the Laser-based R-X filtering is shown in figure 6.

We applied spatial cluster and size filtering to the raw R-X output to produce object detections in both of the R-X image data sets. The spatial clustering and size filtering provides a mechanism to reduce false alarms through the use of both size and shape factors. The resulting detected objects are illustrated in figure 7 for the sections of imagery shown previously. Note that the locations of false alarms and missed detections tend to complement one another in these results. Our JMEX process, described below, is designed to take advantage of this observation.

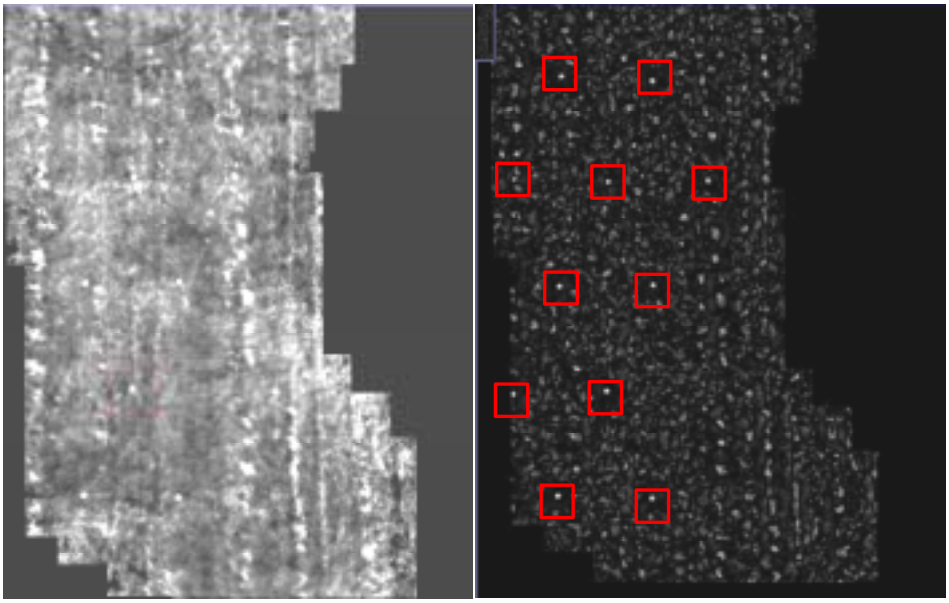


Figure 5. Section of MWIR imagery (left) and R-X filter output (right). Red boxes indicate mine locations.

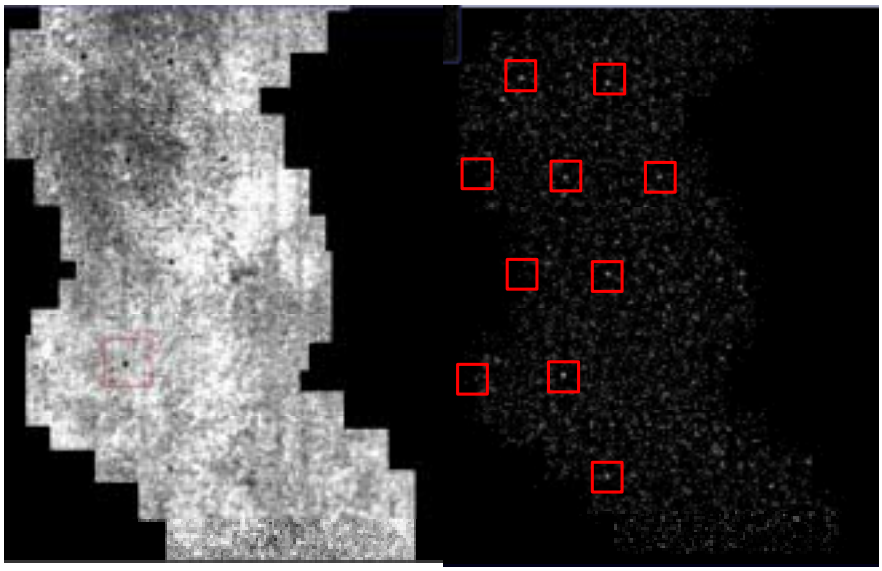
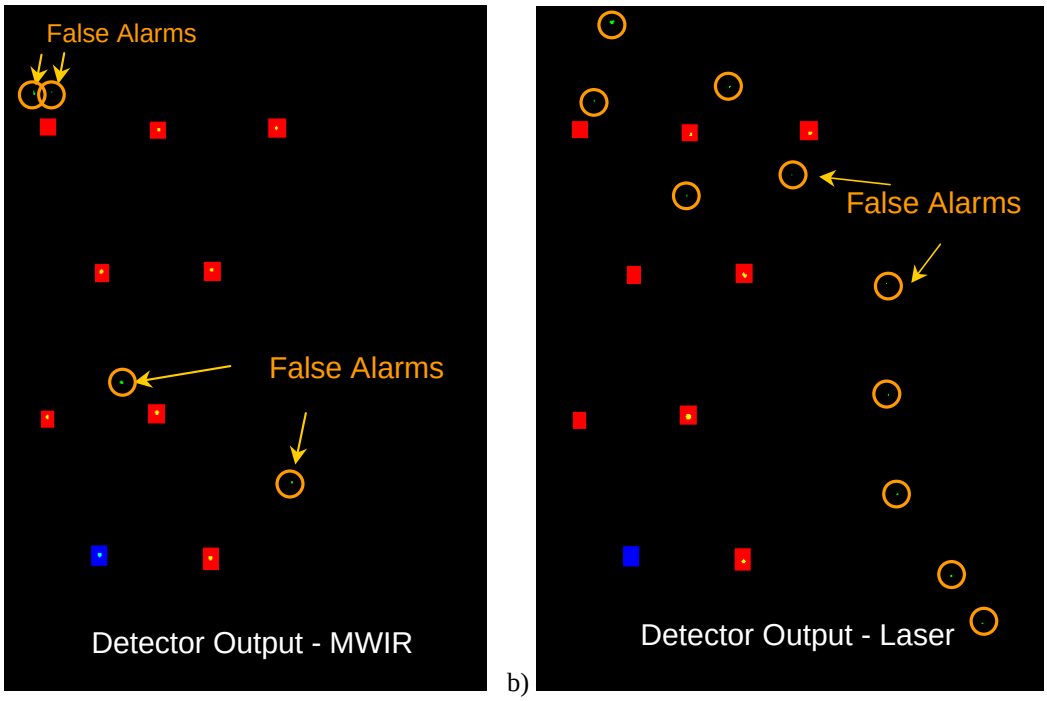


Figure 6. Section of cross-polarized laser-illuminated imagery (left) and R-X filter output (right). Red boxes indicate mine locations.




a) b)

**Figure 7.** Clustered results from R-X applied to a) MWIR imagery and b) Laser imagery, along with target locations in red and blue boxes. Note that the locations of the false alarms and missed detection are different than for the same section of results from the Laser imagery. For illustration the threshold was set at Pd ~ 0.85 over a large segment of this data file containing approximately 20 mine targets.

## 5. JOINT MULTISENSOR EXPLOITATION PROCESSING

Joint Multisensor Exploitation is an adaptation our Joint Algorithm Exploitation (JALEX) approach and is a systematic methodology for fusing registered multiple detector outputs. The method is based on development of a fully adaptive decision criterion from the joint distribution formed by the multiple detector outputs. The JMEX approach seeks to approximate an optimal decision boundary through a process that includes modeling of the marginal probability distributions of each sensor detection feature,

standardizing these marginals and accounting for the correlation between them. Evaluation of the joint distribution for each sample results in a scalar fusion decision statistic such that is thresholded for detection purposes.

### *5.1 Tail Distribution Modeling and Standardization*

The first step in the joint detection feature fusion approach includes modeling the distributions of each detection feature independently (i.e. the marginals of their joint distribution). This is essential for developing joint fusion decision criteria because of the large variability in output distributions from detectors applied to different types of data. Although the detection features previously have many similarities the output distributions may be quite different. The detection feature plane from each sensor is modeled using Gamma mixture distributions to enable standardization of these distributions to a standard distribution, such as exponential or Gaussian as described below.

Quadratic detector distributions may be modeled through the use of Gamma mixture densities based on empirical and theoretical justification [Stein, et al, 2001]. The Gamma mixture model has the following form,

$$\Psi(x) = \sum_{m=1}^{M} w_m \, p_{Gamma}(x|r,a); \sum_{m=1}^{M} w_m = 1 \quad \text{and} \quad w_m \geq 0, \tag{15}$$

where $w_m$; $m$=1,…,$M$ are the mixture weights. The central Gamma probability density function (PDF) with shape parameter $r$ and scale parameter $a$ is given by,

$$p_{Gamma}(x|r,a) = \frac{a^{-r}}{\Gamma(r)} x^{r-1} e^{-x/a}; x > 0. \tag{16}$$

We used a non-linear least-squares fit to the logarithm of the sample tail probability to estimate the parameters of the Gamma mixture to fit each of the marginals. In the figure below we show the result from fitting Gamma mixture models to both the Laser-based and MWIR-based individual detector outputs. In both cases the Gamma mixtures provide and excellent fit to the marginal distributions.

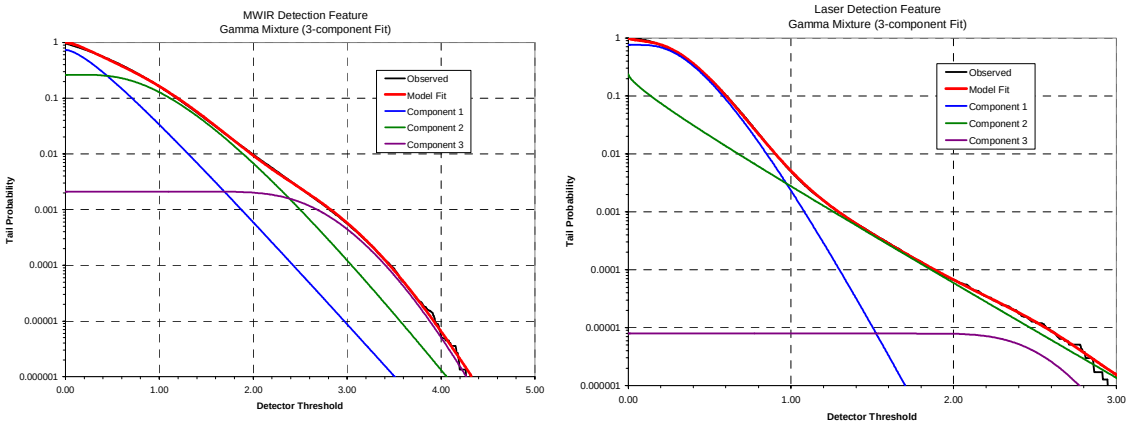



**Figure 8.** Three-component Gamma Mixture model fit to output detector from MWIR imagery (left) and from Laser-illuminated imagery (right).

The Gamma mixture model fit was then used to standardize each marginal distribution to an exponential. The resulting standardization from fitting detection features from two sensors is shown in figure 9. Note that the vastly different distributions shown previously have been effectively standardized to the exponential distribution, which is ideally a straight line in the figure.

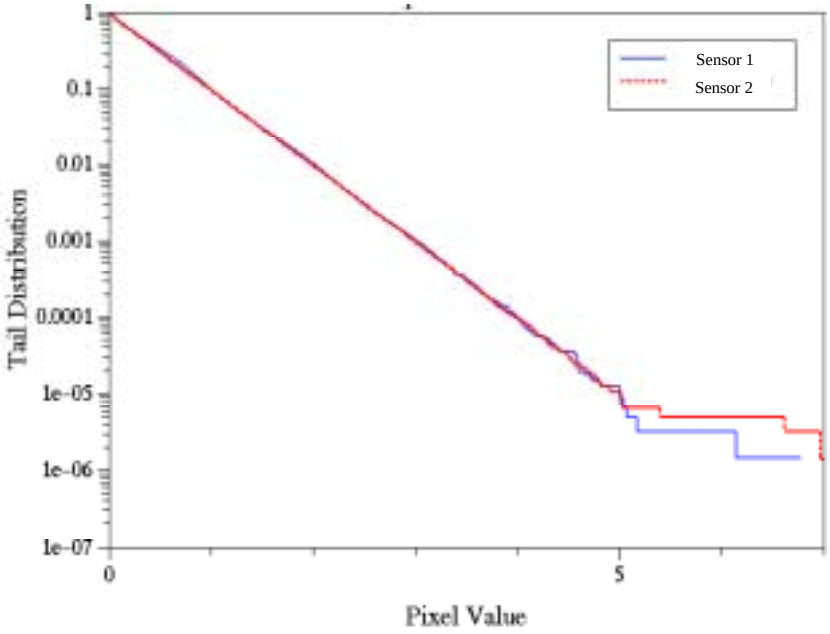



**Figure 19.** Exponential-standardized tail distributions for the multisensor detection. Ideal exponential would be a straight line on this figure. The model fits approximate exponential for five orders of magnitude indicating a good fit between the model and data.

*5.2 Construction of Joint Distribution and Decision Statistics*

To implement joint fusion of detectors from two different sensors (or algorithms) we build a likelihood ratio test based on assumed distributions under the clutter-only ($H_0$) and target ($H_1$) hypotheses. Ideally the distributions are both known and the optimal detector is given by the following

$$\Lambda(\underline{x}) = \frac{f_1(\underline{x}|\underline{\theta}_1)}{f_0(\underline{x}|\underline{\theta}_0)} \underset{H_0}{\overset{H_1}{>}} \gamma, \tag{17}$$

where $\underline{x}$ is the observation vector containing multiple detector output; $f_j$ are the distributions and $\underline{\theta}_j$ are the observations, under hypothesis $j$; and $\gamma$ is a threshold. In our case we do not use any prior target information, which is or equivalent to assuming that the target is uniformly distributed in the detection feature space. However we may construct an estimate for the joint clutter distribution ($H_0$ case) from the sample data and the marginal distributions described in the previous section. Our joint decision rule then becomes a threshold test on the log-likelihood of the data under the joint clutter distribution $f_0$:

$$\lambda(\underline{x}) = -\log\left[ f_0(\underline{x}) \right] \tag{18}$$

This is equivalent to defining a constant-likelihood contour on the clutter distribution in the joint detection feature space. Therefore the construction of a joint distribution is critical in performing the fusion of the multiple sensor outputs. We have two methods for constructing the joint distribution. These are called the "pullback" and "standardized" variants. In the pullback approach the joint density is defined for normal-distributed standardized variables, then "pulled back" to the original feature space of the detector outputs for setting the decision boundary. In the standardized variant the joint density is defined in a transformed feature space of exponential variables, by fitting a multivariate Gamma distribution.

Here we illustrate the concept by application of the standardized variant, using exponential distributions. The construction of the joint distribution is performed in standardized feature space, where the individual detector outputs have been modeled as described in the previous section and standardized to an exponential distribution. The association of samples from one sensor to another may be performed on a pixel level. However, since our expected accuracy of registration between the sensors is greater than a pixel we must

be careful in our association of detector outputs. In this case we associate the largest response within a local window for each of the detectors. This allows for residual misregistration between the sensors. This also allows the individual sensors to respond to different parts of the same object, while still having the detector outputs associated at the object level. The joint distribution is then constructed from the object-level association of the standardized distributions.

The joint distribution is fit with joint Gamma model and this model is fit to the data by estimating the parameters of the marginal Gamma distributions (shape and scale) and the correlation coefficients using least-squares methods. The joint detection statistic is then computed as described above. Figure 10 illustrates the standardized histogram density for the associated samples from the two sensors and model fit (right) for MWIR- and Laser-based detection features. Evaluation of the joint distribution for each sample results in a scalar fusion decision statistic such that is thresholded for detection purposes. Each threshold corresponds to a decision contour in the joint detection feature space shown in figure 10 (right).

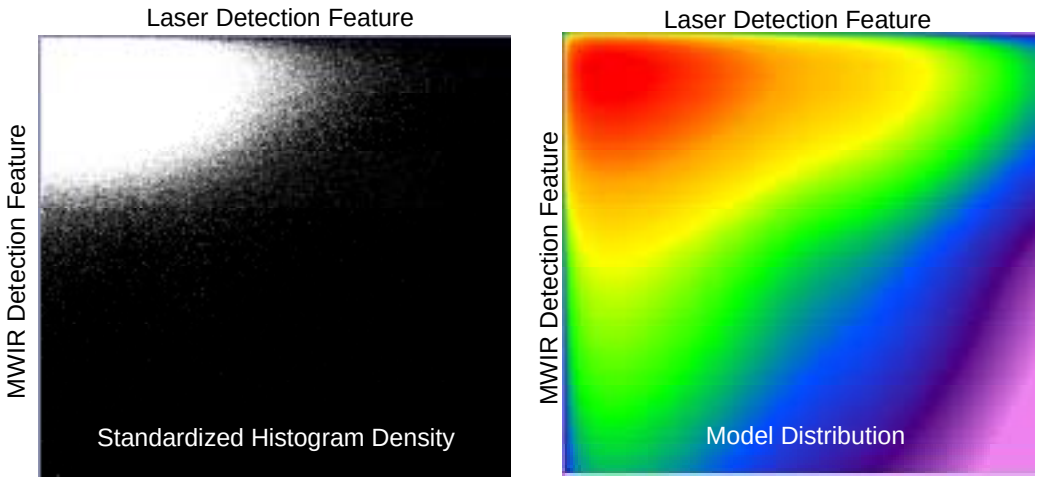



**Figure 10.** Standardized histogram density (left) and standardized joint density model estimate (right), based on the standardized empirical distributions from MWIR- and Laser-based standardized feature space.

## 6. PERFORMANCE RESULTS

The JMEX process, based on local R-X detection features was applied to two different image pairs and scored for detection performance. The scene labeled T281 includes two mine types deployed in tall grass. The T3146 imagery includes two principal mine types deployed on primarily bare soil, with sparse vegetation. Target masks were based on NVESD-provided ground truth text files. We also identified obvious panels in the scenes and ignored responses from these areas in the scoring process. The resulting relative ROC performance metrics for the T281 and T3146 scenes are shown in Figures 11 and 12 below.

In the T281 scene, the better-performing individual sensor was the MWIR camera, whereas in the bare soil case it was the Laser system. In either case, the JMEX processing was as good as the better sensor and in some portions of the ROC curve outperformed detection by either sensor individually by a significant margin. Note that because of the relatively small region care must be taken in extrapolating these performances in general. However, these results indicate that JMEX processing of object-level registered MWIR and Laser-illuminated imagery shows promise in improving overall mine detection performance on surface mines.

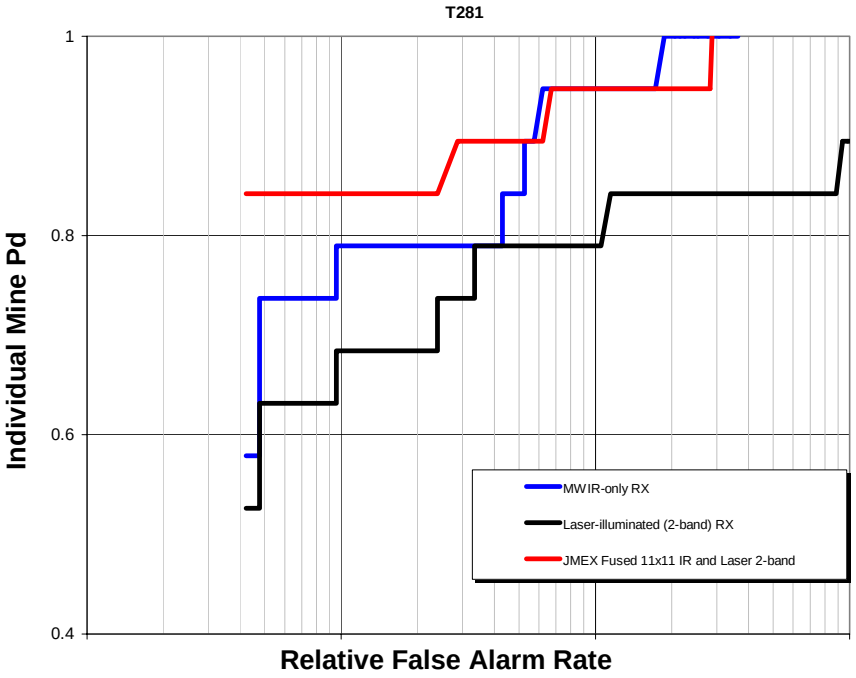



**Figure 11.** Relative ROC performance for the segment of T281 containing all mine targets for individual sensor and JMEX processing. In this example JMEX provided a significant improvement over the better individual sensor (MWIR ).

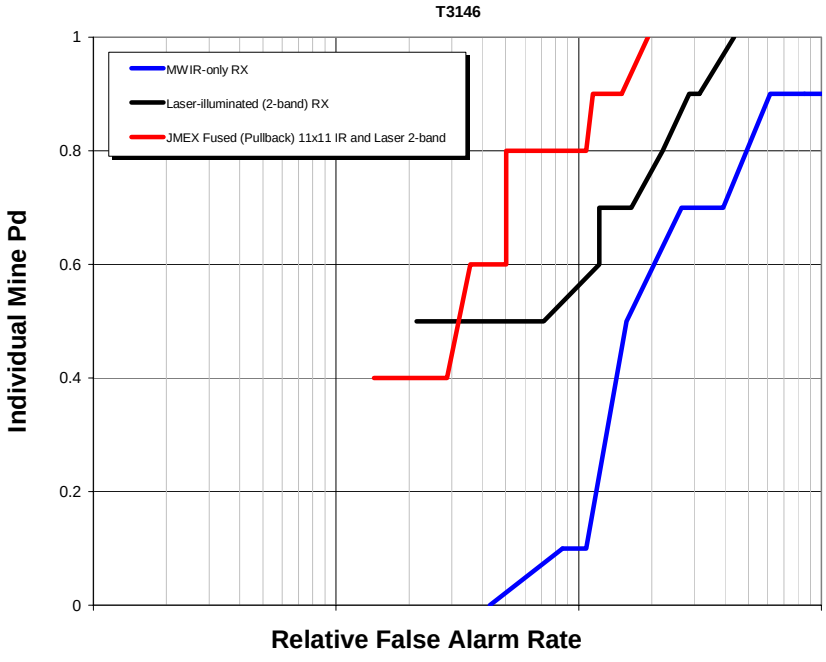



**Figure 12.** Relative ROC performance for the segment of T3146 containing all mine targets for individual sensor and JMEX processing. In this example JMEX provided a significant improvement over the better individual sensor (Laser).

## 7. SUMMARY

In this paper we presented an approach for robust detection of surface mines from multisensor data. The approach is based on joint exploitation of multiple anomaly detection filters and requires neither prior knowledge of target signature nor any training on mine data. The approach was been demonstrated to provide significant detection performance improvement over single sensor airborne mine detection.

## 8. ACKNOWLEDGEMENTS

The authors would like to thank Anh Trang and Chung Phan at NVESD for their help in supplying data as well as feedback on this effort. This effort was performed under a Phase I Army SBIR.

## 9. REFERENCES

1. D.W.J. Stein, S.G. Beaven, L.E. Hoff, E.M. Winter, A.P. Schaum, and A.D. Stocker, "Anomaly Detection from hyperspectral imagery," *Signal Processing Magazine*, Vol. 19, No. 1, Jan 2002.

2. I.S. Reed and X. Yu, "Adaptive Multiband CFAR Detection of an Optical Pattern with Unknown Spectral Distribution," *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. 38, No. 10 (October 1990).

3. H.T. Haskett and R.R. Rupp, "Landmine performance bounds in various background using airborne 808nm laser imagery," Detection and Remediation Technologies for Mines and Minelike Targets VII, Proceedings of the SPIE Vol. 4742 (2002).